

RAW SEQUENCE LISTING

DATE: 09/06/2001

PATENT APPLICATION: US/09/491,146

TIME: 09:10:36

Input Set : N:\Crf3\RULE60\09491146.txt

Output Set: N:\CRF3\09062001\I491146.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: KHUDYAKOV, YURI E.

6 FIELDS, HOWARD A.

8 (ii) TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION

9 ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME

11 (iii) NUMBER OF SEQUENCES: 55

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: JONES & ASKEW, LLP

15 (B) STREET: 191 Peachtree Street, N.W., 37th Floor

16 (C) CITY: Atlanta

17 (D) STATE: GA

18 (E) COUNTRY: USA

19 (F) ZIP: 30303-1769

ENTERED

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/491,146

C--> 29 (B) FILING DATE: 25-Jan-2000

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 08/921,887

34 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: WARREN, WILLIAM L.

39 (B) REGISTRATION NUMBER: 36,714

40 (C) REFERENCE/DOCKET NUMBER: 03063-0380

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 404-818-3700

44 (B) TELEFAX: 404-818-3799

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 55 base pairs

51 (B) TYPE: nucleic acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA

57 (iii) HYPOTHETICAL: NO

59 (iv) ANTI-SENSE: NO

61 (v) FRAGMENT TYPE: internal

63 (vi) ORIGINAL SOURCE:

64 (A) ORGANISM: Hepatitis virus

68 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

70 CCCC GAATTC AACCGAAACC GCAACGTAAA ACCAAACGTA ACACCATTCG TCGTC

55

RAW SEQUENCE LISTING

DATE: 09/06/2001

PATENT APPLICATION: US/09/491,146

TIME: 09:10:36

Input Set : N:\Crf3\RULE60\09491146.txt

Output Set: N:\CRF3\09062001\I491146.raw

72 (2) INFORMATION FOR SEQ ID NO: 2:

74 (i) SEQUENCE CHARACTERISTICS:

75 (A) LENGTH: 69 base pairs

76 (B) TYPE: nucleic acid

77 (C) STRANDEDNESS: single

78 (D) TOPOLOGY: linear

80 (ii) MOLECULE TYPE: cDNA

82 (iii) HYPOTHETICAL: NO

84 (iv) ANTI-SENSE: NO

86 (v) FRAGMENT TYPE: internal

88 (vi) ORIGINAL SOURCE:

89 (A) ORGANISM: Hepatitis virus

93 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

95 CCCC GGATCC TATTTCGGAC CAACGATCTG ACCACCACCC GGAATTAA CGTCCTGCGG 60

97 ACGACGAAT 69

99 (2) INFORMATION FOR SEQ ID NO: 3:

101 (i) SEQUENCE CHARACTERISTICS:

102 (A) LENGTH: 54 base pairs

103 (B) TYPE: nucleic acid

104 (C) STRANDEDNESS: single

105 (D) TOPOLOGY: linear

107 (ii) MOLECULE TYPE: cDNA

109 (iii) HYPOTHETICAL: NO

111 (iv) ANTI-SENSE: NO

113 (v) FRAGMENT TYPE: internal

115 (vi) ORIGINAL SOURCE:

116 (A) ORGANISM: Hepatitis virus

120 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

122 CCCC GAATTC AACCGAAACC GCAACGTCAG ACCAAACGTA ACACCAACCG TCGT 54

124 (2) INFORMATION FOR SEQ ID NO: 4:

126 (i) SEQUENCE CHARACTERISTICS:

127 (A) LENGTH: 70 base pairs

128 (B) TYPE: nucleic acid

129 (C) STRANDEDNESS: single

130 (D) TOPOLOGY: linear

132 (ii) MOLECULE TYPE: cDNA

134 (iii) HYPOTHETICAL: NO

136 (iv) ANTI-SENSE: NO

138 (v) FRAGMENT TYPE: internal

140 (vi) ORIGINAL SOURCE:

141 (A) ORGANISM: Hepatitis virus

145 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

147 CCCC GGATCC TATTTCGGAC CAACGATCTG ACCACCACCC GGAATTAA CGTCCTGCGG 60

149 ACGACGGTTG 70

151 (2) INFORMATION FOR SEQ ID NO: 5:

153 (i) SEQUENCE CHARACTERISTICS:

154 (A) LENGTH: 51 base pairs

155 (B) TYPE: nucleic acid

156 (C) STRANDEDNESS: single

RAW SEQUENCE LISTING

DATE: 09/06/2001

PATENT APPLICATION: US/09/491,146

TIME: 09:10:36

Input Set : N:\Crf3\RULE60\09491146.txt

Output Set: N:\CRF3\09062001\I491146.raw

```

157         (D) TOPOLOGY: linear
159     (ii) MOLECULE TYPE: cDNA
161     (iii) HYPOTHETICAL: NO
163     (iv) ANTI-SENSE: NO
165     (v) FRAGMENT TYPE: internal
167     (vi) ORIGINAL SOURCE:
168         (A) ORGANISM: Hepatitis virus
172     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
174 CCCCGAATTC AACCGAAACC GCAACGTAAA ACCAAACGTA ACACCTACCG T      51
176 (2) INFORMATION FOR SEQ ID NO: 6:
178     (i) SEQUENCE CHARACTERISTICS:
179         (A) LENGTH: 73 base pairs
180         (B) TYPE: nucleic acid
181         (C) STRANDEDNESS: single
182         (D) TOPOLOGY: linear
184     (ii) MOLECULE TYPE: cDNA
186     (iii) HYPOTHETICAL: NO
188     (iv) ANTI-SENSE: NO
190     (v) FRAGMENT TYPE: internal
192     (vi) ORIGINAL SOURCE:
193         (A) ORGANISM: Hepatitis virus
197     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
199 CCCC GGATCC TATTTCCGAC CAACGATCTG ACCACCACCC GGAATTTAA CGTCCTGCGG      60
201 ACGACGGTAG GTG      73
203 (2) INFORMATION FOR SEQ ID NO: 7:
205     (i) SEQUENCE CHARACTERISTICS:
206         (A) LENGTH: 55 base pairs
207         (B) TYPE: nucleic acid
208         (C) STRANDEDNESS: single
209         (D) TOPOLOGY: linear
211     (ii) MOLECULE TYPE: cDNA
213     (iii) HYPOTHETICAL: NO
215     (iv) ANTI-SENSE: NO
217     (v) FRAGMENT TYPE: internal
219     (vi) ORIGINAL SOURCE:
220         (A) ORGANISM: Hepatitis virus
224     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
226 CCCC GAATTC AACCGAAACC GCAACGTAAA CCGAACCGTA ACACCAACCG TCGTC      55
228 (2) INFORMATION FOR SEQ ID NO: 8:
230     (i) SEQUENCE CHARACTERISTICS:
231         (A) LENGTH: 69 base pairs
232         (B) TYPE: nucleic acid
233         (C) STRANDEDNESS: single
234         (D) TOPOLOGY: linear
236     (ii) MOLECULE TYPE: cDNA
238     (iii) HYPOTHETICAL: NO
240     (iv) ANTI-SENSE: NO
242     (v) FRAGMENT TYPE: internal
244     (vi) ORIGINAL SOURCE:

```

RAW SEQUENCE LISTING

DATE: 09/06/2001

PATENT APPLICATION: US/09/491,146

TIME: 09:10:36

Input Set : N:\Crf3\RULE60\09491146.txt

Output Set: N:\CRF3\09062001\I491146.raw

```

245      (A) ORGANISM: Hepatitis virus
249      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
251 CCCCCGATCC TATTTTCGGAC CAACGATCTG ACCACCACCC GGAATTTAA CGTCCTGCGG      60
253 ACGACGGTT      69
255 (2) INFORMATION FOR SEQ ID NO: 9:
257      (i) SEQUENCE CHARACTERISTICS:
258          (A) LENGTH: 63 base pairs
259          (B) TYPE: nucleic acid
260          (C) STRANDEDNESS: single
261          (D) TOPOLOGY: linear
263      (ii) MOLECULE TYPE: cDNA
265      (iii) HYPOTHETICAL: NO
267      (iv) ANTI-SENSE: NO
269      (v) FRAGMENT TYPE: internal
271      (vi) ORIGINAL SOURCE:
272          (A) ORGANISM: Hepatitis virus
276      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
278 CCCCGAATTC AACCGAAACC GCAACGTCAG CCGAAACGTA ACACCCGCG TCGTCCGCAG      60
280 GAC      63
282 (2) INFORMATION FOR SEQ ID NO: 10:
284      (i) SEQUENCE CHARACTERISTICS:
285          (A) LENGTH: 60 base pairs
286          (B) TYPE: nucleic acid
287          (C) STRANDEDNESS: single
288          (D) TOPOLOGY: linear
290      (ii) MOLECULE TYPE: cDNA
292      (iii) HYPOTHETICAL: NO
294      (iv) ANTI-SENSE: NO
296      (v) FRAGMENT TYPE: internal
298      (vi) ORIGINAL SOURCE:
299          (A) ORGANISM: Hepatitis virus
303      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
305 CCCCCGATCC TATTTTCGGAC CAACGATCTG ACCACCACCC GGAATTTAA CGTCCTGCGG      60
308 (2) INFORMATION FOR SEQ ID NO: 11:
310      (i) SEQUENCE CHARACTERISTICS:
311          (A) LENGTH: 55 base pairs
312          (B) TYPE: nucleic acid
313          (C) STRANDEDNESS: single
314          (D) TOPOLOGY: linear
316      (ii) MOLECULE TYPE: cDNA
318      (iii) HYPOTHETICAL: NO
320      (iv) ANTI-SENSE: NO
322      (v) FRAGMENT TYPE: internal
324      (vi) ORIGINAL SOURCE:
325          (A) ORGANISM: Hepatitis virus
329      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
331 CCCCGAATTC AACCGAAACC GCAACGTAAA ACCAAACGTA ACGCTCACCG TCGTC      55
333 (2) INFORMATION FOR SEQ ID NO: 12:
335      (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING

DATE: 09/06/2001

PATENT APPLICATION: US/09/491,146

TIME: 09:10:36

Input Set : N:\Crf3\RULE60\09491146.txt

Output Set: N:\CRF3\09062001\I491146.raw

```

336      (A) LENGTH: 69 base pairs
337      (B) TYPE: nucleic acid
338      (C) STRANDEDNESS: single
339      (D) TOPOLOGY: linear
341      (ii) MOLECULE TYPE: cDNA
343      (iii) HYPOTHETICAL: NO
345      (iv) ANTI-SENSE: NO
347      (v) FRAGMENT TYPE: internal
349      (vi) ORIGINAL SOURCE:
350          (A) ORGANISM: Hepatitis virus
354      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
356 CCCC GGATCC TATTCGGAC CAACGATCTG ACCACCACCC GGAATTAA CGTCCTGCGG      60
358 ACGACGGTG                                                                69
360 (2) INFORMATION FOR SEQ ID NO: 13:
362      (i) SEQUENCE CHARACTERISTICS:
363          (A) LENGTH: 55 base pairs
364          (B) TYPE: nucleic acid
365          (C) STRANDEDNESS: single
366          (D) TOPOLOGY: linear
368      (ii) MOLECULE TYPE: cDNA
370      (iii) HYPOTHETICAL: NO
372      (iv) ANTI-SENSE: NO
374      (v) FRAGMENT TYPE: internal
376      (vi) ORIGINAL SOURCE:
377          (A) ORGANISM: Hepatitis virus
381      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
383 CCCC GAATTC AACCGAAACC GCAACGTAAA AACCAGCGTA ACACCAACCG TCGTC      55
385 (2) INFORMATION FOR SEQ ID NO: 14:
387      (i) SEQUENCE CHARACTERISTICS:
388          (A) LENGTH: 69 base pairs
389          (B) TYPE: nucleic acid
390          (C) STRANDEDNESS: single
391          (D) TOPOLOGY: linear
393      (ii) MOLECULE TYPE: cDNA
395      (iii) HYPOTHETICAL: NO
397      (iv) ANTI-SENSE: NO
399      (v) FRAGMENT TYPE: internal
401      (vi) ORIGINAL SOURCE:
402          (A) ORGANISM: Hepatitis virus
406      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
408 CCCC GGATCC TATTCGGAC CAACGATCTG ACCACCACCC GGAATTAA CGTCCTGCGG      60
410 ACGACGGTT                                                                69
412 (2) INFORMATION FOR SEQ ID NO: 15:
414      (i) SEQUENCE CHARACTERISTICS:
415          (A) LENGTH: 55 base pairs
416          (B) TYPE: nucleic acid
417          (C) STRANDEDNESS: single
418          (D) TOPOLOGY: linear
420      (ii) MOLECULE TYPE: cDNA

```

VERIFICATION SUMMARY

DATE: 09/06/2001

PATENT APPLICATION: US/09/491,146

TIME: 09:10:37

Input Set : N:\Crf3\RULE60\09491146.txt

Output Set: N:\CRF3\09062001\I491146.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:626 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=23
L:653 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=24
L:682 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=25
L:711 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=26
L:740 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=27
L:769 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=28
L:798 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=29
L:827 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=30
L:856 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=31
L:885 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=32
L:914 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=33
L:1425 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=52